IIN THE UNITED STATES DISTRICT COURT
FOR THE WESTERN DISTRICT OF PENNSYLVANIA

UNITED STATES OF AMERICA	)	
	)	
v.	)	Criminal No. 19-369
	)	
LAFON ELLIS		

DEFENDANT'S SUPPLEMENT BRIEF AND RESPONSES TO GOVERNMENT BRIEFS AT ECF NO'S 131 AND 133.

INTRODUCTION

The Defense has filed a Rule 17(c) subpoena for various papers, documents, data, or other objects—specifically TrueAllele's source code. The Defense's need for the source code is to examine/test whether the software program that runs complex mathematical and science-based formulas is verified and validated according to *computer software industry practices* to determine whether the software-generated evidence is foundationally and scientifically reliable and valid under computer science and software engineering principles.

Considering the recent decision disclosing TrueAllele's source code by the New Jersey Court of Appeals, the Defense will first brief relevant portions of the *Pickett* decision before responding to the government's supplements at ECF Nos. 131 and 133.

I. THIS COURT SHOULD FOLLOW THE REASONING OF *NEW JERSEY V. PICKETT* A DECISION WITH IDENTICAL ISSUES, AND EXPERTS, AS THE CASE AT BAR.

1. During the pendency of the source code litigation in Mr. Ellis's case, one other jurisdiction had allowed a defendant access to TrueAllele's source code for independent defense expert examination. *See Virginia v. Watson*, Criminal No. FE-2019-279 (Fairfax Va. Cir. Ct. Oct. 9, 2020) (disclosing TrueAllele source code under protective order).

2. Additionally, since the parties last submitted briefing in this case, the New Jersey Court of Appeals issued an opinion reversing a New Jersey State court judge for failing to disclose

TrueAllele's source code, and related materials, for examination by the defense's experts prior to an admissibility hearing. *See New Jersey v. Pickett*, A-4207-19T4, 2021 WL 357765, at *24 (N.J. Super. App. Div. Feb. 3, 2021).

3. To summarize, two jurisdictions, within the last four months, have recognized that meaningful and independent examination of TrueAllele's source code under software engineering principles is required before determining the admissibility of TrueAllele's software generated evidence. *See Virginia v. Watson*, Criminal No. FE-2019-279 (Fairfax Va. Cir. Ct. Oct. 9, 2020); *New Jersey v. Pickett*, A-4207-19T4, 2021 WL 357765, at *24 (N.J. Super. App. Div. Feb. 3, 2021).

4. Here, the Defense is recommending that the Court follow the *Pickett* Court.

A. THE *PICKETT* COURT HELD THAT TRUEALLELE COMBINES FORENSIC SCIENCE AND SOFTWARE ENGINEERING AND THEREFORE SHOULD BE EVALUATED UNDER BOTH STANDARDS.

5. The *Pickett* Court correctly recognized that "TrueAllele's software integrates multiple scientific disciplines, therefore requiring cross-disciplinary validation to determine reliability." *Pickett*, 2021 WL 357765, at *23.

6. One of the disciplines utilized by TrueAllele to create its software generated evidence is computer science and software engineering principles. *Id*. at *23. The *Pickett* Court was persuaded by the testimony of Drs. Heimdahl and Matthews[1], who explained that "each discipline will validate a program under different standards. In particular, V & V ["Verification and Validation"] in the computer science field cannot be achieved without a thorough examination of the source code which translates validated probabilistic genotyping into executable software." *Id.*

[1] As a reminder to the Court, the Defense hired Drs. Heimdahl and Matthews as experts in this case. Together they co-authored a declaration marked at Exhibit 4 as part of the Defense's brief at ECF No. 121. Dr. Matthews co-authored a declaration with Nathan Adams that will be attached to this brief.

7. The *Pickett* Court emphasized the need for computer science experts to have a seat at the table when determining the foundational validity of software-generated evidence. *Id.* at *23. Emphasizing the need for V & V validation, the *Pickett* Court stated "while TrueAllele may be generally accepted in the field of DNA forensics as methodologically sound, **such validation may be too narrow, thereby making access to the source code <u>even more important to test whether Dr. Perlin's testimony has gained general acceptance in the computer science community to which it also belongs</u>**." *Id.* at *23. While *Pickett* is a *Frye* jurisdiction, this Court must still consider the known or potential error rate, whether <u>the method is generally accepted in the relevant scientific communities</u>, and whether there are standards that control operation and were used in the case at hand pursuant her "gate-keeping functions" under *Daubert v. Merrell Dow Pharmaceuticals, Inc.*, 509 U.S. 579, 593–94 (1993).

8. The *Pickett* Court was also persuaded by the testimony of Nathan Adams,[2] whom they characterized as someone with "important and extensive experience performing probabilistic genotyping analyses, including undertaking review of source codes" of competing probabilistic genotyping systems STRmix and FST "under protective orders in other criminal cases." *Pickett*, 2021 WL 357765, at *5.

9. The *Pickett* court found Mr. Adams's arguments for the need to examine TrueAllele's results under software engineering principles and the "V&V processes" persuasive, writing, "[w]e need not detail every aspect of his declaration; suffice it to say that Mr. Adams provided the judge with an in-depth and thorough basis to grant the motion." *Id.*

[2] As a reminder to the Court, the Defense hired Nathan Adams as an expert in this case. Mr. Adams has written a declaration marked at Exhibit 1 as part of Defense's briefs at ECF Nos. 71, 83. Mr. Adams also wrote a declaration marked at Exhibit 2 as part of ECF No. 121. Mr. Adams co-authored a declaration with Dr. Matthews attached to this brief.

10. In the case at bar, at the request of undersigned counsel, Dr. Matthews and Nathan Adams co-authored a declaration in which they explain the relevance for computer science and software engineering principles in the discussion when evaluating the reliability of TrueAllele. Specifically, they state:

> We emphasize that forensic DNA and computing disciplines both recognize and emphasize the need for validated products and systems, but that the label of "validated" is achieved through different processes in each discipline. We do not intend to diminish the significance or relevance of guidance and standards published by SWGDAM and FBI, ISFG, or ASB, but suggest that **practices common to software development and described in software standards, guidance, articles, and texts are also relevant considerations**. Stated differently, SWGDAM and FBI, ISFG, or ASB do not satisfactorily address issues of software engineering.

See Decl. of Dr. Matthews and Nathan Adams, attached as Exhibit 1 at ¶6. "The source code and executable versions of software programs are highly relevant to any review of verification and validation." Exhibit 1 at ¶9. According to Dr. Matthews and Mr. Adams, the Source code is the human readable list of commands (e.g., "if the user enters X, then do Y," "repeat trying possible solutions until one works") written by programmers and used to generate the executable version itself." *Id*. "The normal user-facing executable version is produced from the source code but is designed to be interpreted by the computer rather than read by a human. When a user runs the executable version of software, it often generates graphics and an interface with which the user can interact." *Id.*

11. Based on the aforementioned reasoning, it is their professional opinion that "[e]xamining both the source code and the executable version are both important and give different windows into potential flaws, including flaws that could affect the likelihood ratio reported by the system." Exhibit 1 at ¶ 9.

12. In their recent declaration Dr. Matthews and Mr. Adams also address Dr. Perlin's claims that source code review is not necessary,

> While Dr. Perlin asserts that access to source code is not necessary to assess the reliability of TrueAllele's results, we simply disagree. The field of software engineering recognizes the importance of source code access and review for common verification and validation tasks. Complex software frequently has latent errors that are difficult to detect and which could impact reported results. Meaningful access to the source code can help uncover those latent defects.

Exhibit 1 at ¶ 14.

B. THE *PICKETT* COURT RECOGNIZED THAT "HIDING THE SOURCE CODE IS NOT THE ANSWER. THE SOLUTION IS PRODUCING IT UNDER A PROTECTIVE ORDER."

13. The *Pickett* Court addressed Dr. Perlin's trade secret arguments, which were mimicked by the state prosecution, as they are by the Government in this case:

> Hiding the source code is not the answer. The solution is producing it under a protective order. Doing so safeguards the company's intellectual property rights and defendant's constitutional liberty interest alike. Intellectual property law aims to prevent business competitors from stealing confidential commercial information in the marketplace; it was never meant to justify concealing relevant information from parties to a criminal prosecution in the context of a Frye hearing.

Pickett, at *1.

14. Relevant to the issues in this case, or ones to arise if this Court denies the Government's motion to quash, the *Pickett* Court also addressed "two key areas of disagreement" between the parties when attempting to come to an agreement about the contents of a protective order. *Pickett*, at *21. "The two areas pertained to liquidated damages for breach of the order, and the terms of the inspection itself." *Id*. The *Pickett* Court left the contents of the protective order to the judge but stated that the judge "should follow these remarks" as to the two issues. *Id*.

15. First, the *Pickett* Court took issue with the State's proposed protective order that asked for "liquidated damages for breach of the order" of "$1,000,000 automatic civil liability 'in the event that the proprietary materials are improperly handled, negligently or otherwise,'" and that "the

defense submit to jurisdiction in Pennsylvania and that the defense obtain liability insurance with $3,000,000 in coverage." *Id*. at *21.

16. As to the liquidated damages issue, the *Pickett* Court noted that "a model protective order from the Northern District of California, whose docket includes among the most complex and financially consequential patent cases in the world, includes no provision for financial liability." *Id*. at *21 (pointing to U.S. Dist. Ct. for the N.D. of Cal., *Model Protective Order for Litigation Involving Patents* ("*Model Protective Order*")[3].

17. The *Pickett* Court concluded its remarks on the liquidated damages provision by stating, "[w]e have not found—and the parties have not provided—any case authorizing disclosure of source code and related proprietary information under a protective order with the restrictions as rigid as Cybergenetics' terms, particularly as to liquidated automatic financial liability for breach of a protective order. *Id*. at *21.

18. Second, the *Pickett* Court addressed "terms of the inspection" of the source code. *Id*. at *22. In *Pickett*, the "State offered to host defense counsel and their experts at the prosecutor's office . . . but then prohibited meaningful inspection by permitting only handwritten notes of 170,000 lines of code . . . ." *Id*. "Moreover, the State required the inspection to be supervised and would not allow photographs or copying of any material." *Id*.

19. The *Pickett* Court again pointed to the model protective order from the Northern District of California, which "includes provisions explicitly permitting certain personnel other than the experts themselves access to the sensitive information, *Model Protective Order* §§ 7.2, 7.3, and

[3] Available at https://www.cand.uscourts.gov/forms/model-protective-orders/ (last visited Feb. 21, 2020).

allows the printing of portions of the source code for purposes of analysis, *id.* § 9(d).” *Pickett* at *22.

20. Dismissing the rigid access to the source code requested by Cybergenetics, the *Pickett* Court found “the defendant’s proposed order to provide[] reasonable protections, including a prohibition on disclosure to any individual with any direct or indirect commercial or employment interest in competing software products.” *Pickett* at *22.

21. As far as meaningful access to the source code, the *Pickett* Court noted that “[a]lthough a requirement that all notes be handwritten may be included to prevent unauthorized copying and disclosure of source code, **such a requirement could be impractical given the form and syntax of source code. Such a requirement may be considered ‘burdensome in the extreme’** because ‘[m]odern computer source code was never intended to be handwritten even by the original programmer.” *Pickett* at *22 (quoting, Lydia Pallas Loren & Andy Johnson-Laird, *Computer Software-Related Litigation: Discovery and the Overly-Protective Order*, 6 Fed. Cts. L. Rev. 1, 47 (2012) (emphasis supplied).

C. IN CHOOSING TO DISCLOSE TRUEALLELE’S SOURCE CODE, THE *PICKETT* COURT WAS GUIDED BY U.S. CONSTITUTIONAL PRINCIPLES.

22. The *Pickett* Court referenced Constitutional principles and United States Supreme Court opinions as a basis for their decision to disclose the source code. *Pickett*, at *13. Focus was given to the criminal defendant’s right to “a meaningful opportunity to present a complete defense.” *Id.* (quoting *Crane v. Kentucky*, 476 U.S. 683, 690 (1986)).

23. “Without access to the source code—the raw materials of the software programming—a defendant’s right to present a complete defense, by meaningful cross-examination at the appropriate juncture, may be substantially compromised.” *Pickett*, at *13. “A criminal trial where the defendant does not have ‘access to the raw materials integral to the building of an effective

defense' is fundamentally unfair." *Id.* (quoting *Ake v. Oklahoma*, 470 U.S. 68, 77 (1985)). "Anything less than full access contravenes fundamental principles of fairness, which indubitably compromises a defendant's right to present a complete defense." *Pickett*, at *24.

D. THE *PICKETT* COURT HELD THAT THE ONLY WAY TO TEST WHETHER THE SOURCE CODE IS PROPERLY IMPLEMENTING THE PROGRAM'S DESIGN SPECIFICATIONS AND THEREFORE PRODUCING RELIABLE RESULTS IS TO TEST IT AND THAT SUCH TESTING SHOULD BE DONE PRIOR TO AN ADMISSIBILITY HEARING.

24. The *Pickett* Court was not persuaded by prior determinations of reliability touted by the prosecution and Dr. Perlin because they recognized that "**prior determinations of reliability in other jurisdictions entailed no scrutiny of computer science or source code**. Instead, the courts depended in large part on Dr. Perlin's own testimony and the existing validation studies which, even if diligently conducted and sound, were **not truly independent and did not even evaluate the source code**. *See Pickett*, at *19.

25. The *Pickett* Court stressed,

> Before going any further, we stress one important point. Evaluating the issues on appeal requires a working knowledge of computer software. Without such a foundation, one can miss subtle consequences germane to this *Frye* hearing. Allowing independent access to the requested information, for the sole purpose of addressing whether the technology underlying the expert testimony is reliable—specifically, whether the source code for that technology is properly implementing the program's design specifications—is obvious. An accused individual's liberty is at stake; DNA evidence is powerful and compelling. **Practically speaking, if, as Dr. Perlin maintains, the source code he wrote is free of harmful defects, and therefore will not impact the reliability of TrueAllele, then it is to everyone's advantage to learn that at the *Frye* hearing. If it should turn out there are source code errors that might affect TrueAllele's reliability, the time to discover that information is now, as part of the judge's gatekeeping role**. Reliability must be resolved at the *Frye* hearing rather than in post-conviction relief proceedings.

Pickett, at *15.

26. The Court concluded by advising courts dealing with similar issues,

> Courts must endeavor to understand new technology—here, probabilistic genotyping—and allow the defense a meaningful opportunity to examine it. **Without scrutinizing its software's source code—a human-made set of instructions that may contain bugs, glitches, and defects—in the context of an adversarial system, no finding that it properly implements the underlying science could realistically be made**. Consequently, affording **meaningful examination** of the source code, which compels the critical independent analysis necessary for a judge to make a threshold determination as to reliability at a *Frye* hearing, **is imperative.**

Pickett, 2021 WL 357765, at *23-24.

E. DISCUSSION.

27. Back to the case at bar, here as in *Pickett*, the same experts are providing their expertise to the Court (Drs. Mats Heimdahl and Jeanna Matthews, and Mr. Nathan Adams). Each wrote declarations explaining the need for disclosure based on computer science and software engineering principles.

28. As Drs. Heimdahl and Matthews said in their November 2020 declaration, none of validation studies and nine "peer reviewed" studies about TrueAllele's reliability, touted by Dr. Perlin and the government, address verification and validation under computer science and software engineering principles. *See* Decl. of Dr. Heimdahl and Dr. Matthews, ECF No. 121-4, re-attached as Exhibit 2 to this filing. They specifically wrote that they

> examined each of the validation studies provided by the prosecution in preparation for this declaration (a total of 39 documents including 9 documents classified as peer reviewed articles). We focused on the 9 peer reviewed articles and concluded the following: 1) The validation studies were not independent, 2) The validation studies did nothing to address the quality of the implementation or likelihood of implementation errors and 3) The validation studies were incomplete. We also point out that peer review of validation studies is simply not the same as software validation and verification. Peer review in a scientific journal indicates only that the reviewers thought the scientific community should see the results, not that the software is reliable enough to be used in the criminal justice system in general or in any specific case in particular.

Exhibit 2 at ¶¶ 41-42.

29. Speaking about Dr. Perlin's declarations, the *Pickett* Court found that "[h]is declaration omits reference to his own involvement in those studies, or the participation in the studies of current or former employees of Cybergenetics, and he neglected to acknowledge the lessons learned from STRmix and FST, which were revealed once other courts forced them to make accessible their source codes for independent review under protective orders." *Pickett* at *6. The *Pickett* Court also took issue with Dr. Perlin not staying in his own lane as an expert in his field—not participating as a legal commentator, finder of fact, or lawyer, something he has done in this case as well—"Dr. Perlin, although he was the State's expert, advocated on behalf of his company that access to the source code would be "immaterial to [a criminal] case," "[un]reasonable," and not "in the interests of justice." *Pickett* at *6.

30. Here, the Defense has argued, *ad nauseum*, for the need to inspect the source code. *See* ECF Nos. 45, 71, 83, and 121.

31. Regarding the need for the source code, this Court seemed to be convinced, and acknowledged, "[d]efendant's good faith, and [that] the broad potential probative relevance of the information sought, are not in doubt." *See Opinion and Order*, ECF No. 132 at 7.

32. In his most recent sworn declaration, Dr. Perlin declared his willingness to disclose the source code under a protective order. In ¶ 57 of his declaration attached as Exhibit 2 at ECF No. 131, Dr. Perlin wrote, "**Cybergenetics makes its TrueAllele source code available to defense teams under confidentiality agreements or protective orders that respect trade secrets**."

33. This Court held in abeyance the decision to disclose the source code as it seemed Dr. Perlin would agree to disclosure under a protective order. *See* ECF No. 132 at 10.

34. When push came to shove however, Dr. Perlin now backtracks his sworn statement stating that "Cybergenetics is not willing to provide its source code absent an order from the Court compelling it to do so." *See* ECF No. 133 at 1.

II. RESPONSES TO SUPPLEMENTS AT ECF NOS. 131 AND 133.

A. RELIABILITY OF THE OUTCOME IS LINKED TO THE SOURCE CODE'S INSTRUCTIONS TO THE SOFTWARE SYSTEM.

35. First, in their filing ECF No. 131, the Government writes "[w]hat the defendant ignores, but what is the key to the analysis, is that the admission of this evidence **TrueAllele depends not on the source code, it is the reliability of the software's outcome**." *See* ECF No. 131 at 1 (emphasis added).

36. The only party ignoring arguments is the Government and Dr. Perlin. Like the Defense continues to argue in this case, the *Pickett* Court also called out Dr. Perlin's refusal to acknowledge that computer science and software engineering principles apply to software generated evidence. *See Pickett* at *6 ("[Dr. Perlin] neglected to acknowledge the lessons learned from STRmix and FST, which were revealed once other courts forced them to make accessible their source codes for independent review under protective orders.").

37. Here, when the Defense argues apples, the government and Dr. Perlin respond with oranges. This tactic misdirects the matter at issue (do computer science and software engineering principles apply to the *Daubert* analysis and can Dr. Perlin prove that TrueAllele follows such principles) by conflating arguments and terms as to what verified and validated and reliable means. Whereas the Defense argues the need for verification and validation under computer science and software engineering principles, Dr. Perlin cites to his "validation studies" and the complex formulae he invented as a reason why this Court should find TrueAllele reliable. Dr. Perlin uses a complex multidiscipline formula that he developed in order to carry out the probabilistic

genotyping analysis. He runs that formula as an algorithm through the secret software code that he developed as well. The software code, authored by Dr. Perlin, implements the probabilistic genotyping algorithm.

38. In an amicus brief filed in *Pickett* by Upturn—an organization dedicated to advances equity and justice in the design, governance, and use of technology—wrote[4]:

> Software can allow more efficient and comprehensive data analysis—but it can also be biased, faulty, or completely ineffective. At the design stage, the process of creating software necessarily includes decisions and assumptions. TrueAllele is no exception. It is these differing design decisions that have resulted in variability in conclusions across probabilistic genotyping software. For example, in a New York case TrueAllele and another probabilistic genotyping software produced different conclusions on the defendant's guilt for the same mixed DNA sample. *See* President's Council of Advisors on Science and Technology (PCAST), Report to the President: Forensic Science in Criminal Courts: Ensuring Scientific Validity of Feature-Comparison Methods 78 n.212 (2016) [hereinafter PCAST Report].[5] This is not a flaw by itself; Cybergenetics should design their own models and write their own code to implement probabilistic genotyping. In fact, these design and programming choices are the precise reason why TrueAllele's developers want to safeguard their code. However, the defense must have access to information about these design choices because they can influence ostensibly objective results. For example, the Forensic Statistical Tool ("FST"), a peer to TrueAllele, was found in a 2016 source code review to have a hidden function that tended to overestimate the likelihood of guilt. *See* Stephanie J. Lacambra et al., *Opening the Black Box: Defendants' Rights to Confront Forensic Software*, NACDL: The Champion (May 2018). Specifically, source code review was able to catch errors in FST. During a murder trial, the court granted a defense expert full access to the program's source code. *See* Lauren Kirchner*, Where Traditional DNA Testing Fails, Algorithms Take Over*, ProPublica (Nov. 4, 2016).[6] This analysis produced two alarming observations. First, the code did not seem to be implementing the methods and models that were used in FST's validation studies. *See* Jessica Goldthwaite et al., *Mixing It Up: Legal Challenges to Probabilistic Genotyping Programs for DNA Mixture Analysis*, Champion (May 2018) at 12, 15 (noting "disturbing differences between what FST was initially advertised to be and what is actually being used in criminal casework"). Second, there seemed to be coding errors that caused results

[4] Upturn was represented by Harvard Law School's Cyberlaw Clinic, and the *amicus* brief was relied on in the *Pickett* Court's decision.

[5] https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/PCAST/pcast_forensic_science_report_final.pdf, [as of Feb. 22, 2020].

[6] https://www.propublica.org/article/where-traditional-dnatesting-fails-algorithms-take-over, [as of Feb. 22, 2020].

> to favor the prosecution's theory of the case. *See id*. Without independent review of TrueAllele's source code, there is no guarantee that TrueAllele does not have similar outcome-determinative functions that may also lead to wrongful convictions and potentially fatal consequences.
>
> Even when software is not designed with faulty assumptions, unintentional errors can significantly impact the software's performance. In 2020, the UK's Most Serious Violence tool, a flagship artificial intelligence system designed to predict future gun and knife violence, was found to have coding flaws that experts concluded made it unusable. Matt Burgess, *Police Built an AI to Predict Violent Crime. It Was Seriously Flawed*, Wired (Aug. 6, 2020).[7] After discovery of a coding error that caused training data to be improperly ingested, the system, originally claimed by its developer to be up to seventy-five percent accurate, was demonstrated to be less than twenty percent accurate. *Ibid*. And in 2015, investigators in Australia encountered an error in their use of STRmix, a probabilistic genotyping software program intended to resolve mixed DNA profiles similar to TrueAllele. David Murray *Queensland Authorities Confirm 'Miscode' Affects DNA Evidence in Criminal Cases*, The Courier Mail (Mar. 20, 2015).[8] The error produced incorrect results in at least sixty criminal cases, including a high-profile murder case. *Ibid*. This is especially concerning given STRmix's striking similarities to TrueAllele—both are forensic identification software systems that use probabilistic genotyping.

Brief of Amicus Curiae Upturn, Inc. in Support of Movant-Appellant, *Pickett*, 2021 WL 357765.[9]

39. The *Pickett* court relied on Upturn's arguments in its decision. *See Pickett*, 2021 WL 357765, at *10 ("Undertaking an independent review establishes whether the software is properly implementing the program's design specifications and that the code itself is devoid of bugs, glitches, and defects that could affect the software's output. And equally important is that TrueAllele's source code has never been scrutinized by any party outside

[7] https://www.wired.co.uk/article/police-violence-predictionndas, [as of Feb. 22, 2020].
[8] https://www.couriermail.com.au/news/queensland/queenslandauthorties-confirm-miscode-affects-dna-evidence-in-criminalcases/news-story/833c580d3f1c59039efd1a2ef55af92b, [as of Feb. 22, 2020].
[9] https://www.upturn.org/static/files/2020-10-15-nj-v-pickett.pdf, [as of Feb. 22, 2020].

of Cybergenetics; therefore, the validation studies produced by the State to date are limited.").

40. Assuming, *arguendo*, that his formulae were scientifically valid—which the Defense does not concede—how can the Defense and Court be sure that the software system that runs his formulae, with his coding, with his assumptions embedded into the program, is not flawed with bugs, errors, and coding mistakes which can affect the likelihood ratio? The answer is unknown, and the Defense and the Court are left with Dr. Perlin's assurances—without proof—and opinion that it is.

41. As the *Pickett* Court said, "[h]iding the source code is not the answer." *Pickett* at *1. Procedural safeguards in later parts of the criminal process afford defendants the opportunity to challenge admitted evidence. *See Daubert*, 509 U.S. at 596 ("Vigorous cross-examination, presentation of contrary evidence, and careful instruction on the burden of proof are the traditional appropriate means of attacking shaky but admissible evidence."). Thus, while maintaining that his software is errorless, Dr. Perlin ignores that competing probabilistic genotyping software programs have been found to be error prone after software inspection which affected the overall conclusion.

42. In stark contrast to Dr. Perlin, the *Pickett* Court emphasized that lessons from competing probabilistic genotyping programs like STRmix and FST, after source code was revealed, "reaffirm the basic principle in computer engineering that software is prone to human error." *Pickett* at *10. According to the *Pickett* Court, "[i]n light of the concerns that arise when examining the "black box" validation studies, the out-of-state judicial opinions and orders that have accepted TrueAllele's reliability without source code examination, and errors found in the source codes of the breathalyzer in *Chun*, FST, and STRmix, judges should examine the reliability of such software with healthy skepticism." *Pickett*, at*20. Thus, contrary to what the government asserts, the

reliability of the software's outcome needs to be tested from all the different disciplines involved in the production of the final outcome—the likelihood ratio. Andrea Roth, *Machine Testimony*, 126 Yale L.J. 1972, 1982 (2017)[10] ("Courts applying *Daubert-Frye* to software-generated statements should treat software engineers as part of the relevant scientific community and determine reliability not only of the method, but also of the software implementing that method, based on industry standards. In some cases, courts would likely need to access proprietary source code to assess the code's ability to operationalize an otherwise reliable method."). The fact that Dr. Perlin is keeping the source code under lock and key invites even more uncertainty.

B. THE SCALE ANALOGY IS FLAWED, AND SUBSEQUENT TESTING DEMONSTRATES THE NEED TO EXAMINE TRUEALLELE'S SOFTWARE.

43. The government's scale analogy, and subsequent testing of Mr. Ellis's DNA using competing "open source" probabilistic genotyping programs is troubling, *see* ECF No. 131 at 1-2, and seemingly unbeknownst to them further emphasizes Defense's need for access to the source code. In their analogy, the government attempts to analogize to the measurement of a scale, arguing that just like the accuracy of a scale can be measured by weighing the same substance on different scales, the accuracy of TrueAllele can be measured by testing the results with different "open source" probabilistic genotyping systems. *Id*. at 1-2. According to the government, preventing access to the source code is necessary because "opening up the scale and examining the inner workings would provide no information as to accuracy of the scale and may destroy the scale." *Id.* at 1. The government goes on by writing, "[w]e need not examine the source code of TrueAllele to determine whether it works. Rather, we can test the results TrueAllele produces to determine whether it produces reliable results, just as we would with a scale." *Id*. at 1-2.

[10] https://www.yalelawjournal.org/pdf/RothFinal_c4o97on1.pdf

44. When slightly extended, the scale analogy quickly falls apart. One "thing" (i.e., the similarity of Mr. Ellis's DNA to the evidence vs. the similarity of the general population to the evidence) was weighed by Cybernetics on six "scales" (i.e., probabilistic genotyping software programs) with the question posed, "Does this thing weigh more than 1 pound or less than 1 pound?" While all scales said, "more than 1 pound," each scale came up with a different absolute weight. One scale says the thing weighs 1,400 pounds – similar to the weight of a horse. Another scale said the thing weighs 1 quadrillion pounds – the weight of Lake Erie and Lake Ontario, which is 10% of the total Great Lakes water.[11] Is "horse" concordant with "lake"?

45. Here, the Defense is not just concerned with the "more than 1 pound" conclusion. We are also concerned with the "*how much* more than 1 pound" conclusion. Moreover, looking at six other similar-ish programs does not tell us that the program—TrueAllele—works, particularly in light of the disparity in results it produced. And giving access to the source under a protective order will not "break" TrueAllele, as the government analogized.

46. Looking at other programs does not tell us how TrueAllele reached its conclusion, which is the central question at issue here—TrueAllele's results are what will be introduced as evidence by the government at trial. Additionally, the fact that other programs that purport to do the same analysis as TrueAllele actually produce different results is concerning. A recent study [hereinafter "Garofano study"] comparing the results of different probabilistic genotyping software programs illustrates how different programs find different answers. Paolo Garofano, *et al.,* Forensic Science International: Genetics Supplement Series 5, *An alternative application of the consensus method to DNA typing interpretation for Low Template-DNA mixtures* (2015) pp. 422-424. The study uses

[11] https://www.britannica.com/topic/Areas-and-Volumes-of-the-Great-Lakes-1800353, [as of Feb. 22, 2020].

three different probabilistic genotyping programs to analyze five samples taken from different evidence items and comprising mixtures of multiple contributors.

47. For one sample, two of the three programs calculated inconclusive likelihood ratios of 1.20 and 1.29. The third program, however, reported an inclusionary statistic of 109 trillion. For a second set of samples, two programs again reported exclusionary likelihood ratios in the hundreds—arguably in an inconclusive range. The third program, however, reported an inclusionary likelihood ratio in the hundred millions. For a third item, all three programs reported inclusionary likelihood ratios: 900 million, 1 billion or 5 hundred quintillion. The greatest likelihood ratio was a trillion times larger than the smallest likelihood ratio.

48. The Garofano study examined only a few samples and got answers from different probabilistic genotyping software programs that were sometimes strikingly divergent. The study, however, did not propose an explanation for why this occurred. Is one program miscalculating? Are there types of samples where one of the programs becomes less reliable? Because likelihood ratios lack an objectively measurable ground truth, it is not possible to discern from a program's results alone whether accurate, false, or wildly misleading results are being produced. Source code review is necessary to determine whether the program being utilized in casework is actually implementing the methodology described in publications, and whether the assumptions it incorporates are sound.

49. In another study, the California Department of Justice compared the results of mock samples run through TrueAllele with the results of the same mock samples run through a competing software program STRmix. In one part of the study, TrueAllele resulted in false inclusions in 18% of a set of samples where STRmix resulted in none. In a second set of samples STRmix resulted in false inclusions in 3.7% of the cases and TrueAllele in none. A third set of

samples found false inclusion rates of 3.4% for STRmix and 5.1% for TrueAllele. *See* California Department of Justice, *NicheVision Sole Source Justification STRmix Mixture Interpretation Software* (Aug. 25, 2014) p. 3-4.[12]

50. The most publicized discordant results between probabilistic genotyping programs occurred in the New York case of *People v. Hillary*, St. Lawrence Cty. Ind. No. 2015-15 (Aug. 26, 2016) (Catena, J.). There, TrueAllele results for a crime scene sample produced inconclusive results, while its commercial competitor STRmix produced inculpatory results. The STRmix results were precluded by the court because the lab that had performed the testing had never validated STRmix.

51. The *Hillary* case brought to the national spotlight the need for full disclosure of the functions that created the final results, including the source code. In fact, the case was specifically referenced in an addendum to the recently released report by the PCAST Report:

> A recent controversy has highlighted issues with [probabilistic genotyping (PG)]. In a prominent murder case in upstate New York, a judge ruled in late August (a few days before the approval of PCAST's report) that testimony based on PG was inadmissible owing to insufficient validity testing. Two PG software packages (STRMix and TrueAllele), from two competing firms, reached opposite conclusions about whether a DNA sample in the case contained a tiny contribution (~1%) from the defendant. Disagreements between the firms have grown following the conclusion of the case. (Citations omitted).

52. The discordant results in *Hillary* between TrueAllele and STRmix have not been addressed by the scientific community, except by the commercial developers of these programs themselves who, without offering any specifics, each claim that their software program modeled the data better than the other.

[12] Available at, https://epic.org/state-policy/foia/dna-software/EPIC-16-02-02-CalDOJ-FOIA-20160219-Procurement-Justification-STRmix.pdf, [as of Feb. 22, 2020].

53. The assumptions inherent in a program's design may help explain the discordant results. The scientific community is keenly aware of the effect such assumptions have on probabilistic genotyping results. *See* Hinda Haned, et.al., *Validation of probabilistic genotyping software for use in forensic DNA casework: Definitions and illustrations* (2016) Science and Justice 56, pp. 104–108 ("Model validation for use in forensic casework is not straightforward because the true weight of the DNA evidence cannot be determined; indeed, the generated [likelihood ratio] always depends on the model's assumptions, no 'gold standard' exists in the form of a true likelihood ratio that can serve as a comparison."). No scientist, let alone any lawyer, can discern how various assumptions embedded in the source code affect the likelihood ratio without access to those assumptions. In light of the wide range of potential outcomes seen in the Garofano study and elsewhere, it is particularly problematic for a court to effectively preclude cross-examination as to the fundamental design choices found in the source code.

C. THE DEFENSE IS ACTING IN GOOD FAITH IN THIS LITIGATION AND INTENDS TO REVIEW THE SOURCE CODE ONCE DISCLOSED.

54. The government appears to be arguing that merely litigating substantive challenges to the reliability of their evidence amounts to bad faith. *See* ECF No. 131 at 3. But as the Court has recognized "[d]efendant's good faith, and [that] the broad potential probative relevance of the information sought, are not in doubt." *See Opinion and Order*, ECF No. 132 at 7.

55. In contrast, some of the government's recent filings contain mischaracterizations of similar litigation in another jurisdiction, specifically the *Virginia v. Watson* case. *See* ECF No. 131 at 4; *See als*o Decl. of Mark Perlin ¶ 60, ECF No. 131-2. Bryan Kennedy, a Senior Assistant Public Defender at the Office of the Public Defender, in Fairfax County, Virginia, who is counsel of record in the Watson case, wrote a declaration to this Court detailing the "multiple misstatements

of fact," made by Dr. Perlin and the government in this case. *See* Decl. of Bryan Kennedy, attached as Exhibit 3.

56. As to the assertion that the Defense is not interested in reviewing the source code once disclosed, once provided with meaningful access, the Defense experts stand ready to perform the sort of analysis that is necessary to assess the reliability of Cybergenetics's software, which is necessary to assist the Court in her gate-keeping function under *Daubert*.

57. As the Mr. Adams and Dr. Matthews detail, what they need access to is the source code and executable version of the software program. *See* Exhibit 1 at ¶9 ("The source code and executable versions of software programs are highly relevant to any review of verification and validation."). Meaningful access to test the reliability of the program is not accomplished by simply going over to Cybergenetics's office and running the program as Dr. Perlin and the government keep suggesting.

D. THE PROTECTIVE ORDER IN *VIRGINIA V. WATSON* WILL NOT ALLOW MEANINGFUL ACCESS, DR. PERLIN KNOWS THAT, AND THAT IS WHY HE IS REQUESTING THE SAME ORDER HERE.

58. In the event this Court where to disclose the source code, the government, and Dr. Perlin are asking for a protective order akin to the one issued in *Virginia v. Watson*, Criminal No.: FE-2019-279 (Fairfax Va. Cir. Ct. October 9, 2020). *See* ECF No. 133 at 2.

59. Contrary to the government's assertion, Mr. Kennedy went to look at the source code and in his declaration to this Court, described the format of the production. *See* Exhibit 3 at ¶ 12.

60. Mr. Kennedy's description of the format in which the source code was produced, does not constitute meaningful access as described to be minimally necessary by Dr. Matthews and Mr. Adams, in their declaration to this Court. According to them,

> At a minimum, meaningful access to examine TrueAllele's source code requires the code to be on a personal computer, with a full keyboard and mouse. The code

> needs to be "buildable" into a working executable from the MATLAB IDE with all necessary libraries and working build instructions provided. In addition, the ability to install code analysis tools is necessary to evaluate the code.

See Exhibit 2 at ¶ 19. Meaningful review requires "access to relevant tools . . . with which the program and its development processes can be measured, tested, and evaluated from a variety of perspectives." *Id*. at ¶ 15.

III. CONCLUSION

Based on the foregoing and all the briefs and declarations presented, this Court should deny the motion to quash the subpoena for TrueAllele's source code, and enter a protective order consistent with the recommendation of Pickett, and that guarantees the Defense meaningful access to the code. The Defense has drafted a proposed protective order that we believe would achieve meaningful access, while adequately protecting Cybergenetics interests.

Respectfully Submitted,

/s/ *Khasha Attaran*
Khasha Attaran
Assistant Federal Public Defender